

GISELA HARRAS/KRISTEL PROOST

STRATEGIEN DER LEMMATISIERUNG VON IDIOMEN

Abstract

In diesem Beitrag beschäftigen wir uns mit der Frage der Lemmatisierung von Idiomen in allgemeinen einsprachigen, alphabetisch geordneten, Wörterbüchern des Deutschen. Die drei wichtigsten Fragen der Lemmatisierung von Idiomen sind die folgenden:

1. Unter welchem Stichwort oder Lemma sollte ein Idiom auftauchen?
2. An welcher Stelle des Wörterbuchartikels sollte ein Idiom auftauchen?
3. Unter welcher Nennform sollte ein Idiom auftauchen?

Wir plädieren dafür, diese Fragen in Übereinstimmung mit vier allgemeinen Lemmatisierungsprinzipien zu beantworten und geben dann einen Überblick über die Behandlung von Idiomen in der gegenwärtigen lexikographischen Praxis, wobei wir deren Vor- und Nachteile erläutern. Abschließend formulieren wir drei Maximen für die Lemmatisierung von Idiomen und zeigen, wie diese befolgt werden können, ohne dass es zu einem Bruch mit der gegenwärtigen Praxis kommt.

This paper deals with the lemmatisation of semi- and non-literal idioms in monolingual general dictionaries of German which are alphabetically ordered. The following issues are crucial to the question of how idioms should be lemmatised:

1. Under which headword should an idiom be entered in a general dictionary?
2. Which position should a given idiom occupy within the entry of the headword under which it is listed?
3. What is the appropriate citation form in which the idiom should be entered?

We argue that these questions should be answered in accordance with a set of general lemmatisation principles. We show how these issues have been dealt with in some of the most widely known German dictionaries and discuss the advantages and disadvantages of current lexicographic practice. To avoid lexicographic inconsistencies, we suggest a set of maxims for the lemmatisation of idioms. These provide practicable solutions to each of the basic questions mentioned above.

1. Einleitung

In diesem Beitrag geht es um die Einordnung von Idiomen in den Lemmabestand eines Wörterbuchs. Seit den achtziger Jahren gibt es zahlreiche metalexikographische Überlegungen zur Lemmatisierung von Idiomen (vgl. z.B. Braasch 1988, Burger 1983, Burger 1998, Schemann 1991). Unseres Erachtens wird in all diesen Überlegungen der lexikalische Status von Idiomen zu wenig berücksichtigt. Mit diesem Beitrag wollen wir deshalb versuchen, diese Lücke zu schließen. Wir werden uns im Wesentlichen auf allgemeine einsprachige Wörterbücher des Deutschen mit alphabetischer bzw. nestalphabetischer Anordnung der Lemmata beschränken. Bevor wir auf die einzelnen Aspekte der Lemmatisierung eingehen, werden wir kurz sagen, welche Wortverbindungen wir zu den Idiomen zählen und wie diese sich von anderen lexikalisierten Wortverbindungen unterscheiden.

2. Der lexikalische Status von Idiomem

2.1 Phraseologismen

Idiome sind eine Klasse von Wortverbindungen, die in der Literatur meist als phraseologische, feste oder usuelle Wortverbindungen oder auch als Phraseologismen bezeichnet werden. Wir verwenden den Begriff „Phraseologismen“ als Oberbegriff für alle Arten solcher Wortverbindungen. Phraseologismen unterscheiden sich von Wörtern bzw. von freien Wortverbindungen durch die Eigenschaften der Polylexikalität einerseits und der Stabilität andererseits (vgl. Baranov/Dobrovol'skij 1999, S. 64; Burger 1998, S. 14; Fleischer 1997, S. 36-37; Dobrovol'skij 1995, S. 14-17). Phraseologismen sind komplexe Lexikalisierungen. Durch die Eigenschaft der Polylexikalität unterscheiden sie sich von Wörtern, die einfache Lexikalisierungen darstellen. Die Anwendung grammatischer Regeln ist bei Phraseologismen stärker restringiert als bei freien Wortverbindungen. Manche Verben und Nomen, die Teil einer phraseologischen Wortverbindung sind, kommen nur in bestimmten Flexionsformen vor (vgl. Burger 1998, S. 22-23; Dobrovol'skij 1995, S. 38-39). Einige der in den Phraseologismen vorkommenden Verben sind auf ein bestimmtes Tempus festgelegt, vgl.:

einen Narren an jmdm. gefressen haben

→ *Er frisst einen Narren an seiner Nachbarin.

→ *Er fraß einen Narren an seiner Nachbarin.

→ *Er wird einen Narren an seiner Nachbarin fressen.

in jmdn. ist der Teufel gefahren

→ *In ihn fährt der Teufel.

→ *In ihn fuhr der Teufel.

→ *In ihn wird der Teufel fahren.

Nomen, die Teil eines Phraseologismus ausmachen, kommen häufig nur im Singular oder Plural vor, vgl.:¹

in den sauren Apfel beißen

→ *Da werden wir wohl in die sauren Äpfel beißen müssen.

Etwas geht jmdm. auf den Keks.

→ ?Das geht mir gewaltig auf die Kekse.

sich mit fremden Federn schmücken

→ Das Verdienst war gar nicht sein eigenes; *er hat sich nur mit einer fremden Feder geschmückt.

¹ (In den Beispielen bedeutet „#“, dass die intendierte – hier die idiomatiche – Lesart nicht zustande kommt. „?“ bedeutet, dass der Satz weder in seiner idiomatichen noch in seiner wörtlichen Lesart wohlgeformt ist.):

Viele Phraseologismen erlauben ein gewisses Maß an Variation ihres Konstituentenbestandes, aber die Möglichkeiten, einen Ausdruck (oder mehrere Ausdrücke) eines Phraseologismus durch andere Ausdrücke mit ähnlicher Bedeutung zu ersetzen, sind begrenzt. Wenn ein Wort eines Phraseologismus durch ein anderes, bedeutungsähnliches Wort ersetzt wird, verliert die Wortverbindung oft ihre phraseologische Bedeutung und wird als freie Wortverbindung interpretiert (vgl. Burger 1998, S. 23), vgl.:

jmdm. einen Floh/ *eine Fliege/ *eine Mücke ins Ohr/ *in die Nase/ *hinters Ohr setzen

Die Möglichkeiten der lexikalischen Substitution, die ein bestimmter Phraseologismus erlaubt, können nicht auf Grund von Selektionsrestriktionen erklärt werden, da diese sich auf ganze Paradigmen von Wörtern wie z.B. semantische Klassen oder Felder beziehen. Die Restriktionen, denen die Kookkurrenz von Wörtern in festen Wortverbindungen unterliegt, betreffen hingegen nur einzelne Elemente solcher Paradigmen, vgl.:

jmdm. Honig/ Brei/ *Marmelade um den Bart/ ums Maul/ um den Mund/ *um die Lippen schmieren

[Warum *Honig* und *Brei*, aber nicht *Marmelade*; Warum *um den Bart*, *ums Maul* und *um den Mund*, aber nicht *um die Lippen*?]

Obwohl die meisten Phraseologismen die Anwendung bestimmter syntaktischer Operationen erlauben, sind sie in syntaktischer Hinsicht insgesamt weniger flexibel als freie Wortverbindungen. Phraseologismen bilden hinsichtlich ihrer syntaktischen Variationsmöglichkeiten keine in sich homogene Klasse: Manche erlauben ein breites Spektrum an syntaktischen Variationen und Modifikationen, während andere relativ invariabel sind (vgl. Dobrovolskij 2000; Fraser 1970; Nunberg/Sag/Wasow 1994). Insgesamt erlauben Phraseologismen nur eine bestimmte Anzahl derjenigen syntaktischen Operationen, die bei freien Wortverbindungen grundsätzlich möglich sind. Die syntaktischen Restriktionen der Flexibilität von Phraseologismen sind zwar auch aus lexikographischer Sicht interessant, haben aber keine unmittelbaren Konsequenzen für die Auswahl der Nennform eines Phraseologismus oder für die Art und Weise der Einordnung in den Lemmabestand eines Wörterbuchs.

2.2 Kollokationen und Idiome

Polylexikalität und strukturelle Stabilität sind Eigenschaften aller Phraseologismen. Darüber hinaus weisen manche Phraseologismen ein gewisses Maß an Idiomatizität auf. Unter Idiomatizität verstehen wir die Nicht-Kompositionalität der Bedeutung: Eine feste Wortverbindung ist idiomatisch, wenn ihre Bedeutung sich nicht aus den usuellen, d.h. wörtlichen oder lexikalisch-metaphorischen Bedeutungen ihrer Bestandteile zusammensetzt. Einzelne Typen phraseologischer Wortverbindungen unterscheiden sich hinsichtlich ihres Idiomatizitätsgrades. Idiome haben insgesamt einen höheren Idiomatizitätsgrad als andere Typen fester Wortverbindungen (vgl. Dobrovolskij 1995, S. 19-20). Wir verwenden den Ausdruck *Kollokationen* zur Bezugnahme auf alle Typen von festen Wortverbindungen, die kaum oder gar nicht idiomatisch sind. Beispiele solcher nicht-

oder schwach-idiomatischen Phraseologismen sind Funktionsverbgefüge wie *ein Versprechen geben* oder *Maßnahmen treffen* sowie komparative Phraseologismen, die völlig transparent sind, wie *hart wie Stein*, *schwarz wie die Nacht* usw. Die Grenze zwischen der Klasse der Kollokationen und der der Idiome ist fließend.

Idiome bilden keine in sich homogene Klasse: Sie sind nicht alle gleich idiomatisch. Manche enthalten Komponenten, die ihre freie Bedeutung (d.h. die Bedeutung, die sie außerhalb des Idioms haben) beibehalten. Solche Idiome nennen wir teil-idiomatische Idiome oder Teil-Idiome. Idiome, die keine Komponenten mit freier Bedeutung enthalten, sind „voll-idiomatisch“ (vgl. Burger 1998, S. 32). Das Vorhandensein von Komponenten mit freier Bedeutung erhöht den Kompositionalitätsgrad eines Idioms. Teil-Idiome sind daher immer kompositional im Fregeschen Sinn. Dennoch können auch manche voll-idiomatischen Idiome als kompositional betrachtet werden. Vollidiomatische Idiome gelten als kompositional, wenn zur Erklärung ihrer Bedeutung eine Paraphrase angegeben werden kann, in der Teile ihrer lexikalischen Struktur mit Teilen ihrer Bedeutung korrespondieren. Nach Nunberg, Sag and Wasow wird diese Art der Kompositionalität „Dekomponierbarkeit“ oder „Teilbarkeit“ genannt (vgl. Nunberg/Sag/Wasow 1994). Ein Beispiel für ein teilbares Idiom ist *den Wald vor lauter Bäumen nicht sehen*. Zur Erklärung der Bedeutung dieses Idioms kann eine Paraphrase gefunden werden, in der Teile der formalen Struktur des Idioms mit Teilen seiner Bedeutung übereinstimmen. Der Komponente *den Wald* kann die Bedeutung ‚das große Ganze‘, der Komponente *Bäumen* die Bedeutung ‚Einzelheiten‘ und *sehen* die Bedeutung ‚erkennen‘ zugeordnet werden. Ein Idiom wie *ins Gras beißen* weist keine solche Homomorphie von Form und Bedeutung auf: Zur Erklärung der Bedeutung dieses Idioms kann keine Paraphrase gefunden werden, in der Teile der formalen Struktur des Idioms mit Teilen seiner Bedeutung übereinstimmen. Dementsprechend gilt es als nicht-teilbar. (Vgl. Abb. 1) Idiome wie *den Wald vor lauter Bäumen nicht sehen* sind teilbar, weil sie paraphrasierend in Teile zerlegt werden können. Dies bedeutet aber nicht, dass sich Sprecher beim Produzieren und Verstehen von Idiomen einer solchen Teilbarkeit bewusst wären. Mit Teilbarkeit eines Idioms ist gemeint, dass Sprecher grundsätzlich in der Lage sind, solche Idiome in sinnvolle Bestandteile zu zerlegen (z.B. wenn sie nach der Bedeutung eines Idioms gefragt werden). Diese Fähigkeit ist Teil der Kompetenz eines Sprechers einer bestimmten Sprache. Bei der Teilbarkeit eines Idioms handelt es sich also um eine Kompositionalität post factum (vgl. Nunberg/Sag/Wasow 1994, S. 496; Dobrovol'skij 2000b, S. 115). Mit dem Kompositionalitätsprinzip Freges hat diese Art von Kompositionalität nichts zu tun: Teilbare voll-idiomatische Idiome sind immer nicht-kompositional, denn ihre Bedeutung setzt sich nicht aus den usuellen Bedeutungen ihrer Bestandteile zusammen. Insgesamt ergibt sich also für die Klasse der Phraseologismen die folgende Einteilung:

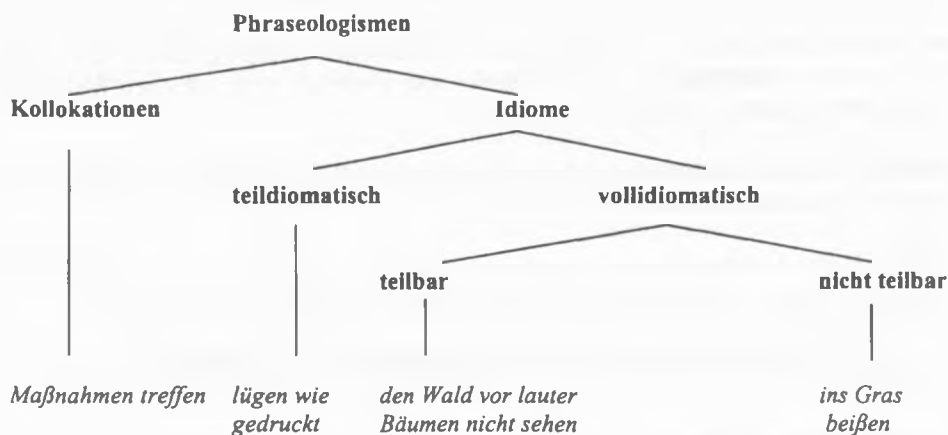


Abb.1: Klassifikation von Phraseologismen

Strukturelle Stabilität, Idiomatizität und Teilbarkeit bestimmen den Grad der Kompositionalität und somit den lexikalischen Status einer fester Wortverbindung. Wir werden uns in diesem Beitrag auf die Lemmatisierung von teil- und voll-idiomatischen Ausdrücken beschränken.

3. Die Lemmatisierung von Idiomen: Drei grundsätzliche Fragen

Angesichts dieses Befundes der lexikalischen Struktur von teil- und vollidiomatischen Ausdrücken stellen sich für ihre Lemmatisierung drei grundsätzliche Fragen (vgl. auch Braasch 1988, Burger 1983, S. 57-62; Burger 1998, S. 170-178; Schemann 1989, S. 2790-2792):

1. Die Frage nach der makrostrukturellen Einordnung, d.h.: Unter welchem Stichwort oder Lemma soll ein Idiom eingeordnet werden?
2. Die Frage nach der mikrostrukturellen Einordnung, d.h.: An welcher Stelle eines Artikels soll ein Idiom aufgeführt werden? Insbesondere muss hier die Frage beantwortet werden, ob es sinnvoll ist, ein Idiom einem der Bedeutungspunkte eines bestimmten Lemmas zuzuordnen.
3. Die Frage nach der Nennform, d.h.: In welcher Nennform soll ein Idiom aufgeführt werden?

Die Beantwortung dieser drei Fragen sollte durch die folgenden vier Prinzipien geleitet sein:

Prinzip 1: Die Lemmatisierung sollte in der Weise erfolgen, dass die Idiome in ihrer Eigenschaft als Elemente der Makrostruktur eines Wörterbuchs repräsentiert sind, d.h. die Ausdrücke müssen, da sie den Status lexikalischer Einheiten haben, als eigene (Sub)Lemmata erkennbar werden. Sie müssen also graphisch in der gleichen Weise ausgezeichnet sein wie alle anderen Lemmata auch.

Prinzip 2: Die Lemmatisierung sollte in Abhängigkeit von der Struktur und dem lexikalischen Status eines jeweiligen Idioms erfolgen.

Prinzip 3: Die Lemmatisierung sollte keine zusätzlichen semantischen Probleme schaffen (wie z.B. eine unnötige Vermehrung der Anzahl der Bedeutungspunkte).

Prinzip 4: Die Lemmatisierung sollte mit der aktuellen lexikographischen Praxis vereinbar sein.

Um das vierte Prinzip befolgen zu können, bedarf es einer Einschätzung der gängigen lexikographischen Praxis. Wir werden uns daher im Folgenden mit den makro- und mikrostrukturellen Einordnungsgepflogenheiten der Wörterbuchmacher sowie ihren Praktiken der Nennformauswahl beschäftigen.

4. Die lexikographische Praxis

4.1 Die lexikographische Tradition der makro-strukturellen Einordnung

In der lexikographischen Praxis haben sich für die der Lemmatisierung von Idiomen im wesentlichen die folgenden zwei Traditionen etabliert:

- (1) Idiome werden unter jedem Inhaltswort, das Bestandteil des Idioms ist, aufgeführt und erklärt.

Dieses Verfahren hat die folgenden Vorteile:

- Die Idiome sind leicht auffindbar.
- Alle Idiomvarianten sind unter der Komponente, die ihnen gemeinsam ist, vollständig aufgeführt:

frieren wie ein Schneider/ wie ein junger Hund

WDG: aufgeführt und erklärt unter: **Schneider**: frieren wie ein Schneider

Hund: frieren wie ein junger Hund

frieren: frieren wie ein Schneider/ wie ein junger Hund

Da alle Idiomvarianten unter mindestens einer der Idiomkomponenten vollständig aufgelistet sind, wird dieses Lemmatisierungsverfahren der lexikalischen Flexibilität der betreffenden Idiome gerecht.

Allerdings hat das Verfahren auch die folgenden Nachteile:

- Die mehrfache Lemmatisierung von Idiomen beansprucht viel Platz und ist daher nur für umfangreiche Printwörterbücher und elektronische Wörterbücher geeignet.
- Das Verfahren ist sehr fehleranfällig. Da Idiome bei jeder ihrer Komponenten aufgeführt werden müssen, kann es leicht passieren, dass die Erwähnung eines Idioms bei einer seiner Komponenten vergessen wird. Wenn das Verfahren konsequent eingehalten würde, müssten die Idiome an jeder Stelle in der gleichen Nennform und mit dem gleichen Bedeutungskommentar erscheinen, was aber häufig nicht der Fall ist:

eine Stange angeben

WDG: unter **angeben** (3. Bedeutungspunkt: ‚prahlen, aufschneiden‘): SALOPP

Der gibt aber eine Stange an! [kein Bedeutungskommentar]

unter **Stange** (1. Bedeutungspunkt: ‚langer, dünner, gerader Gegenstand ...‘):

eine Stange angeben; Bedeutung: ‚prahlen, großsprecherisch auftreten‘
keine Stilangabe.

- (2) Idiome werden unter dem ersten im Idiom auftretenden Nomen aufgeführt und erklärt. Wenn das Idiom kein Nomen enthält, erscheint es unter dem ersten „semantisch signifikanten“ Wort (meist ein Inhaltswort). Querverweise bei den anderen Idiomkomponenten leiten die Benutzer zu dem Lemma, unter dem das jeweilige Idiom erklärt ist. In manchen Wörterbüchern, wie z.B. im HWDG, wird systematisch von Querverweisen Gebrauch gemacht; in anderen erscheinen Querverweise eher sporadisch oder gar nicht.

Dieses Verfahren hat die folgenden Vorteile:

- Es ist platzsparend und übersichtlich, vor allem wenn auf Querverweise verzichtet wird, und selbst wenn systematisch von diesen Gebrauch gemacht wird, müssen die Idiome nur an einer Stelle erklärt werden.
- Da Bedeutungskommentare und Stilangaben nur einmal aufgeführt werden, ist das Verfahren nicht besonders fehleranfällig.

Es hat aber auch die folgenden Nachteile:

- Die Konvention, Idiome nur unter dem ersten in ihnen auftretenden Nomen zu erklären, wird der Struktur und der lexikalischen Variabilität der betreffenden Idiome oft nicht gerecht:

DUDEN & DUW

frieren wie ein Schneider

nur unter **Schneider**; nicht unter **frieren**

frieren wie ein junger Hund

nur unter **Hund**; nicht unter **frieren**

DUDEN & DUW

lügen wie gedruckt (unter **lügen**)
lügen, dass sich die Balken biegen (unter **Balken**)
jmdm. die Hücke voll lügen (unter **Hücke**)
jmdm. die Jacke voll lügen (unter **Jacke**)

Problematisch ist, dass hier auf die Möglichkeit verzichtet wird, alle Idiomvarianten an einer Stelle zusammenzuführen, was die Struktur dieser Teil-Idiome besser reflektieren würde. So bleibt unklar, dass es sich überhaupt um Varianten eines einzelnen Teil-Idioms handelt. Die *frieren*-Idiome werden gar nicht unter der stabilen Idiomkomponente aufgeführt; von den *lügen*-Idiomen erscheint nur eins unter der Komponente *lügen*. Dieses Problem könnte zumindest teilweise durch einen systematischen Gebrauch von Querverweisen behoben werden wie in den folgenden Beispielen:

HWDG

frieren wie ein Schneider
erklärt unter **Schneider** – Querverweis bei **frieren**

HWDG

lügen wie gedruckt (erklärt unter **lügen** – Querverweise auf **Balken**, **Hücke**, und **Jacke**)
lügen, dass sich die Balken biegen (erklärt unter **Balken** – Querverweis auf **lügen**)
jmdm. die Hücke voll lügen (erklärt unter **Hücke** – Querverweis auf **lügen**)
jmdm. die Jacke voll lügen (erklärt unter **Jacke** – Querverweis auf **lügen**)

- Das Kriterium, ein Idiom nur unter dem ersten Nomen oder sonstigem Inhaltswort aufzuführen, scheint willkürlich, was man auch daran erkennen kann, dass es nicht immer konsequent angewendet ist. Besonders wenn ein Idiom kein Nomen enthält, ist oft nicht klar, unter welcher seiner Komponenten es erklärt werden soll:

DUDEN & DUW

sich schwarz/ grün und blau/ gelb und grün ärgern (unter **ärgern**)
aber:
jmdn. grün und blau/ grün und gelb schlagen (unter **grün**)
(Hier werden zwei Idiome mit identischem lexikalischem Status und lexikalischen Varianten unterschiedlich behandelt.)

- Für den Benutzer ist das Kriterium nicht immer leicht nachvollziehbar: *angeben wie eine Tüte Mücken* dürfte eher unter *angeben* als unter *Tüte* nachgeschlagen werden. Solche Überlegungen werden offensichtlich auch von einigen Lexikographen berücksichtigt, vgl.:

DUDEN/OXFORD

angeben wie eine Tüte Mücken
unter **angeben**
(aber: frieren wie ein Schneider unter **Schneider** usw.)

Die Auszeichnung der Idiome als Elemente der Makrostruktur, die wir in unserem Prinzip 1 gefordert haben, ist in den gängigen Wörterbüchern insgesamt sehr uneinheitlich: Manchmal sind die Idiome fett gedruckt, d.h. als eigene Lemmata ausgezeichnet, manchmal sind sie aber auch kursiv gesetzt, so dass ein naiver Benutzer auf die Idee kommen könnte, dass die verzeichneten Ausdrücke lediglich Beispiele präferierter Verwendungsweisen darstellen.

4. Die lexikographische Tradition der mikro-strukturellen Einordnung

In Hinblick auf die Frage, welche Stelle Idiome innerhalb eines Wörterbuchartikels einnehmen sollten, haben sich ebenfalls zwei Traditionen etabliert:

Idiome werden einem der Bedeutungspunkte des Lemmas zugeordnet, vgl.:

DUDEN & DUW

jmdm. einen Bären aufbinden

Unter **Bär**; 1. großes Raubtier (2.: Rammklotz, großer Hammer)

Die Zuordnung von Idiomen zu den einzelnen Bedeutungspunkten eines Lemmas wird häufig damit begründet, dass sie das Verständnis der Idiome und ihre Speicherung im Gedächtnis erleichtern würde (vgl. Scholze-Stubenrecht 1988, S. 293). Diese Begründung beruht wohl auf die Annahme, dass die Zuordnung eines Idioms zu den Bedeutungspunkten eines Lemmas auf der Folie der Entstehungsgeschichte des betreffenden Idioms geschieht. Wenn auf etymologische Erklärungen verzichtet wird, bleibt aber unklar, wie die wörtliche Bedeutung der Idiomkomponente, unter der das betreffende Idiom eingeordnet wird, mit ihrer idiomatischen Bedeutung zusammenhängt. Wenn beispielsweise das Idiom *jmdm. einen Bären aufbinden* unter dem ersten Bedeutungspunkt des Nomens *Bär* (,großes Raubtier') eingeordnet wird, wird aus der Zuordnung nicht deutlich, was die wörtliche Bedeutung von *Bär* mit der idiomatischen Bedeutung dieser Idiomkomponente (,Lüge' oder ,Lügengeschichte') gemeinsam hat. Da dieser Zusammenhang durch eine solche Zuordnung allein nicht erkennbar wird, ist es fragwürdig anzunehmen, dass sie zum Verständnis und zur Einprägsamkeit des betreffenden Idioms beitragen könnte.

Oftmals wird die Praxis der Zuordnung von Idiomen zu den einzelnen Bedeutungspunkten eines Lemmas mit dem Argument gerechtfertigt, dass die meisten Idiomkomponenten semantische Merkmale mit den entsprechenden lemmatisierten Wörtern gemeinsam hätten (vgl. Scholze-Stubenrecht 1988, S. 293-294). Diese Begründung ist aber wenig stichhaltig, weil die idiomatische Bedeutung der Komponenten vieler voll-idiomatischer Idiome gerade keine Gemeinsamkeiten mit ihrer wörtlichen Bedeutung aufweist. So ist beispielsweise unklar, was die wörtliche Bedeutung der Komponente ,Bär' (,großes Raubtier') mit ihrer idiomatischen Bedeutung ,Lüge' gemeinsam hat. Ähnliches gilt für die wörtliche Bedeutung der Komponente ,Floh' (,kleines, blutsaugendes Insekt') und deren idiomatische Bedeutung, ,unerfüllbarer Wunsch' in *jmdm. einen Floh ins Ohr setzen*.

Die Zuordnung von Idiomen zu Bedeutungspunkten eines Lemmas kann zu einer künstlichen Remotivierung völlig opaker Idiome führen (vgl. Burger 1998, S. 175). Um die Bedeutung eines Idioms adäquat wiedergeben zu können, werden oftmals zusätzliche Bedeutungspunkte eingeführt. Dadurch werden Wörtern Lesarten zugeschrieben, die sie außerhalb des betreffenden Idioms nicht haben, vgl.:

DUDEN & DUW

jmdm. einen Korb geben

unter **Korb** 4.: eine ablehnende Antwort auf ein Angebot, einen [Heirats]Antrag

Dieses Beispiel unterscheidet sich von den vorherigen Beispielen dadurch, dass eine Idiomkomponente nicht mit ihrer freien, sondern mit ihrer phraseologisch gebundenen Bedeutung assoziiert wird. Diese Assoziation ist deswegen möglich, weil das Idiom teilbar ist. Für das Idiom *jmdm. einen Korb geben* kann eine Paraphrase gefunden werden, in der der Komponente *Korb* die Bedeutung ‚eine ablehnende Antwort auf ein Angebot‘ zugeordnet werden kann. Dies bedeutet aber nicht, dass *Korb* diese Bedeutung auch außerhalb des Idioms hat.

- (2) Die zweite Tradition der mikrostrukturellen Einordnung besteht darin, dass Idiome am Ende des Artikels, d.h. unterhalb aller Bedeutungspunkte eines Lemmas, aufgeführt werden. Vorangestellte graphische Zeichen machen den Status der Wortverbindung als Idiom oder als Phraseologismus klar, vgl.:

HWDG

jmdm. einen Korb geben

unter **Korb**

1. geflochtener fester Behälter

2. Korbgeflecht

3. (Sport) Ring, an dem ein offenes Netz befestigt ist, in das der Ball beim Korbball, Basketball geworfen werden muss.

4. Handschutz am Griff von Degen, Säbeln

+ jmdm. einen Korb geben, sich einen Korb holen, einen Korb bekommen

4.3 Die lexikographische Tradition der Nennform-Angabe von Idiomen

Statt der Nennform werden oft aktualisierte Beispiele aufgeführt, d.h. Formulierungen, in denen entweder nur eine spezielle flektierte Verbform (anstatt des Infinitivs) auftritt oder die Argumentstellen des Idioms durch ganz bestimmte Nominalausdrücke besetzt sind. Die Angabe spezieller Formen ist allerdings in vielen Fällen irreführend, weil sie zu starke Beschränkungen signalisiert, vgl.:

WDG

Jmd. lügt das Blaue vom Himmel herunter.

Das ist Wasser auf seine Mühle.

Da die Nennform von verbalen Idiomen entsprechend der Nennform von Verben im Infinitiv angegeben wird, sind Restriktionen hinsichtlich der Besetzung der Subjektstelle oft nicht aus der Nennform ersichtlich, vgl.:

DUDEN & DUW

Wasser auf jmds. Mühle sein.

→ Etwas/ *Jemand ist Wasser auf jmds. Mühle.

Andererseits werden durch die Nennform manchmal auch zu starke Restriktionen suggeriert, vgl.:

WDG

einen Streit/ Krieg vom Zaun brechen

DUDEN & DUW

einen Streit/ Krieg/ Zwist o.ä. vom Zaun brechen

HWDG

einen Streit vom Zaun/ Zaune brechen

Tatsächlich sind mehr Varianten erlaubt als diese Nennformen suggerieren, vgl.:

einen Streit/ ein Streitgespräch/ eine Debatte/ eine Diskussion/ ein Gespräch vom Zaun brechen.

Aber: *eine Rede/ *eine Gardinenpredigt vom Zaun brechen

In diesem Fall scheint es möglich, die Substitutionsmöglichkeiten aufgrund von Selektionsbeschränkungen zu erklären: X vom Zaun brechen, wobei X = eine kommunikative Handlung, an der mehr als eine Person (Partei) beteiligt ist. Die Komponente X wäre dann eine idiom-externe Komponente.

5. Maximen zur Lemmatisierung von Idiomen

Zur Lösung der drei grundsätzlichen Fragen, die sich bei der Lemmatisierung von Idiomen stellen, schlagen wir die drei folgenden Maximen vor.

5.1 Die Maxime der makro-strukturellen Einordnung

Durch die Einordnung eines Idioms in den Lemmabestand eines bestimmten Wörterbuchs muss dessen lexikalischer Status deutlich werden. Aus dieser Maxime folgt, dass teil- und voll-idiomatische Ausdrücke hinsichtlich ihrer Lemmatisierung unterschiedlich behandelt werden müssen: Teil-idiomatische Ausdrücke werden unter ihrer wörtlichen Komponente aufgeführt und erklärt, Querverweise bei den übrigen Idiomkomponenten geben dem Benutzer einen Hinweis auf den Erklärungsort, vgl.:

jmdm. den Himmel auf Erden versprechen (erklärt unter **versprechen**; bei **Himmel**: Verweis auf **versprechen**)

jmdn. über den grünen Klee loben (erklärt unter **loben**; bei **Klee/ grün**: Verweis auf **loben**)

angeben wie eine Tüte Mücken (erklärt unter **angeben**; bei **Tüte/Mücken**: Verweis auf **angeben**)

jmds. Lob singen (erklärt unter **Lob**; bei **singen**: Verweis auf **Lob**)

frieren wie ein Schneider (erklärt unter **frieren**; bei **Schneider**: Verweis auf **frieren**)

dumm wie Bohnenstroh (erklärt unter **dumm**; bei **Bohnenstroh**: Verweis auf **dumm**)

Voll-idiomatische Ausdrücke werden unter dem ersten in ihnen auftretenden Nomen aufgeführt und erklärt. Wenn das Idiom kein Nomen enthält, wird es unter dem ersten Inhaltswort aufgeführt und erklärt. Bei den übrigen Idiomkomponenten wird das Idiom nur aufgeführt und mit einem Verweis auf die Komponente versehen, unter der es erklärt wird. Voll-idiomatische Ausdrücke werden also im Gegensatz zu teil-idiomatischen Ausdrücken rein formal behandelt, vgl.:

jmdm. einen Bären aufbinden (erklärt unter **Bär**, bei **aufbinden**: Verweis auf **Bär**)

zwei Fliegen mit einer Klappe schlagen (erklärt unter **Fliege**; bei **Klappe/schlagen**: Verweis auf **Fliege**)

das Kind mit dem Bade ausschütten (erklärt unter **Kind**; bei **Bad/ ausschütten**: Verweis auf **Kind**)

Sowohl teildiomatische als auch vollidiomatische Ausdrücke müssen so ausgezeichnet werden, dass für den Benutzer erkennbar wird, dass sie zum Lemmabestand des Wörterbuchs gehören, d.h. Elemente der Makrostruktur sind.

Wir haben uns dafür entschieden, vollidiomatische Ausdrücke unter dem ersten in ihm auftretenden Nomen einzuordnen, weil dieses Verfahren mehr als alle anderen an die aktuelle lexikographische Tradition anschließt. Da das Kriterium für die Einordnung der voll-idiomatischen Ausdrücke rein formaler Natur ist, wäre irgendeine andere rein formale Lösung ebenso denkbar. So wäre es beispielsweise auch möglich, das Idiom als ein eigenes Lemma alphabetisch unter seinem ersten Wort einzuordnen, unabhängig davon, ob das erste Wort ein Inhalts- oder Funktionswort ist. Nach diesem Verfahren würde ein Idiom wie *jmdm. einen Bären aufbinden* unter Vernachlässigung seiner Argumente als ein eigenes Lemma und in der Form *einen Bären aufbinden* alphabetisch unter dem Buchstaben „e“ eingeordnet werden. Ein Idiom wie *ins Gras beißen* würde unter dem Buchstaben „i“ als eigenes Lemma angesetzt werden usw. Da voll-idiomatische Ausdrücke ohnehin nicht nach strukturellen oder inhaltlichen Gesichtspunkten eingeordnet werden, wäre dieses Verfahren strikt formal und daher auch konsequenter als die Einordnung dieser Ausdrücke unter dem ersten in ihnen auftretende Nomen. Allerdings würde ein solches Verfahren auch zu einer Anhäufung von Idiomem an gewissen Stellen des Wörterbuchs führen. Unter „e“ würden beispielsweise zunächst alle Idiome, die mit dem unbestimmten Artikel anfangen, erscheinen. Diese müssten dann alphabetisch ge-

ordnet werden, d.h. es würden zuerst die Idiome, die mit [*ein* + N] anfangen, erscheinen, dann diejenigen mit [*eine* + N] und schließlich diejenigen, die mit [*einen* + N] anfangen. Zur letzten Gruppe würden etwa Idiome wie *einen Bären aufbinden*, *einen Korb geben* und *einen Floh ins Ohr setzen* gehören. Damit diese Idiome aufgefunden werden können, müssten sie in einer bestimmten Reihenfolge aufgelistet werden. Dazu werden die in ihnen auftretenden Nomen alphabetisch angeordnet:

1. einen Bären aufbinden
2. einen Floh ins Ohr setzen
3. einen Korb geben

Die konsequente Einordnung unter dem ersten Wort führt also durch die Hintertür das Prinzip der Einordnung unter dem ersten Nomen doch wieder ein. Der einzige Unterschied zu der von uns vorgeschlagenen Strategie besteht darin, dass bei diesem Verfahren die Einordnung unter dem ersten Nomen ein zweites untergeordnetes Prinzip ist, während sie bei dem von uns vorgeschlagenen Verfahren das einzige Prinzip der makrostrukturellen Einordnung darstellt.

Wenn das rein formale Verfahren, nach dem voll-idiomatische Ausdrücke als eigene Lemmata alphabetisch unter ihrem ersten Wort eingeordnet werden, nicht ohne das Kriterium des ersten Nomens auskommt, ist es ökonomischer, voll-idiomatische Ausdrücke wie gewohnt unter ihrem ersten Nomen oder sonstigen Inhaltswort anzusetzen. Außerdem schließt die Einordnung unter dem ersten Nomen an die lexikalische Tradition an, was die Benutzerfreundlichkeit der entsprechenden Wörterbücher erhöht. Die Einordnung von Idiomen unter ihrem ersten Inhalts- oder Funktionswort würde das vierte Lemmatisierungsprinzip verletzen, wonach die Lemmatisierung von Idiomen nicht zu einem Bruch mit der lexikographischen Praxis führen darf.

Entscheidend für die Einordnung von Idiomen in allgemeinen alphabetischen und nest-alphabetischen Wörterbüchern ist das Vorhandensein oder Nicht-Vorhandensein von Idiomkomponenten mit freier Bedeutung. Dieses Kriterium ist sowohl für Lexikographen als auch für Wörterbuchbenutzer transparent und daher leicht anzuwenden. Für Lexikographen dürfte es außerdem nicht allzu schwer sein, das Kriterium einheitlich einzuhalten. Für den Benutzer ist festzuhalten, dass er als Person zu antizipieren ist, die ein **Idiom** nachschlägt – und nicht etwa ein isoliertes Wort. Das heißt: der Benutzer muss die formale Struktur des Idioms durchschaut haben, um überhaupt sinnvoll zum Wörterbuch greifen zu können. Oder, um mit Wittgenstein zu sprechen: Wer nach der Bedeutung eines Wortes in einer Sprache fragt, muss schon eine ganze Menge über diese Sprache wissen.

Mit dem Kriterium ‚teil- vs. voll-idiomatisch‘ kann eine große Menge idiomatischer Ausdrücke einheitlich behandelt werden: Es kann sowohl auf alle phraseologischen Vergleiche der Form [*V wie/dass* XP] oder [*A wie* NP] als auch auf phraseologische Verbindungen [*A+N*] angewendet werden, vgl.:

sich schwarz /grün und blau/ grün und gelb ärgern (alle unter **ärgern** erklärt)
 angeben wie eine Tüte Mücken/ wie ein Wald voll Affen/ wie zehn nackte Ne-
 ger/ eine Stange angeben (alle unter **angeben** erklärt)
 dumm wie Bohnenstroh (unter **dumm** erklärt)
 der kalte Krieg (unter **Krieg** erklärt, weil voll-idiomatisch; oder unter **kalt**,
 weil *Krieg* eine lexikalisierte Metapher für ‚Konflikt‘ ist)
 blinder Passagier (unter **Passagier** erklärt weil eine Erweiterung der wörtlichen
 Bedeutung von *Passagier*)
 blutiger Anfänger (unter **Anfänger**, weil teil-idiomatisch)
 blutiger Ernst (unter **Ernst**, weil teil-idiomatisch)

Insgesamt kann man sagen, dass das Verfahren den lexikalischen Status eines Teil-Idioms optimal verdeutlicht, weil alle Varianten eines Teil-Idioms an einer Stelle zusammengeführt und da auch erklärt werden. Dadurch wird das Strukturmuster phraseologischer Vergleiche adäquat repräsentiert. Das Verfahren ist ökonomisch: Ein einziger Bedeutungskommentar genügt zur Erklärung der Bedeutung mehrerer Teil-Idiome mit gemeinsamer wörtlicher Komponente (z.B. *sich schwarz/grün und blau/ grün und gelb ärgern*: ‚sich sehr ärgern‘). Auf diese Weise wird auch die gemeinsame Bedeutung dieser Idiome hervorgehoben.

Dadurch, dass voll-idiomatische Ausdrücke im Gegensatz zu den Teil-Idiomen rein formal behandelt werden, gibt es für die Lemmatisierung von Idiomen zwei Einordnungsprinzipien anstelle von nur einem. Da die Lemmatisierung von Teil-Idiomen nun aber die gleiche ist wie die von Kollokationen, die ebenfalls unter der wörtlichen Komponente lemmatisiert werden, erhöht sich die Anzahl der Lemmatisierungsprinzipien für Phraseologismen insgesamt nicht.

5.2 Die Maxime der mikro-strukturellen Einordnung

Diese Maxime ist für Teil-Idiome irrelevant: Teil-Idiome werden unter der Komponente mit wörtlicher Bedeutung eingeordnet. Voll-idiomatische Ausdrücke werden am Ende des Artikels, d.h. unterhalb aller Bedeutungspunkte des Lemmas, unter dem sie eingeordnet sind, aufgeführt. Dadurch werden Schwierigkeiten und Willkür bei der Zuordnung von Idiomen zu einzelnen Bedeutungspunkten vermieden.

5.3 Die Maxime der Nennform-Angabe von Idiomen

Grundsätzlich sollte ein Idiom in seiner Grundform oder kanonischen Form aufgeführt werden. Die Nennform sollte daher nur allgemeine Pronomina wie *jemand* und *etwas* enthalten, und verbale Idiome sollten grundsätzlich im Infinitiv des im Idiom auftretenden Verbs angegeben werden. In einigen Fällen werden allerdings Abweichungen von der kanonischen Nennform erforderlich, wobei nach den folgenden Untermaximen verfahren werden sollte:

- Die Nennform muss so restriktiv wie möglich sein.

Im Gegensatz zur Valenz von Wörtern ist die Valenz von Idiomen idiosynkratisch. Die Valenz von *schlagen* ergibt sich beispielsweise aus der lexikalischen Bedeutung von *schlagen* (*Jmd. schlägt jmdn.*), aber die Valenz des Idioms *Etwas schlägt jmdm. auf den Magen* ergibt sich nicht aus der freien Bedeutung von *schlagen* (**Jmd. schlägt jmdm. auf den Magen*). Da Restriktionen hinsichtlich der Besetzung der Argumentstellen von Idiomen idiosynkratisch sind, müssen sie im Wörterbuch angegeben werden. Obligatorische Beschränkungen der Besetzung der Subjektstelle müssen aus der Nennform hervorgehen, vgl.:

Etwas hält sich die Waage. (nicht: sich die Waage halten)

Etwas ist ein Spiel mit dem Feuer. (nicht: ein Spiel mit dem Feuer sein)

Etwas schlägt jemandem auf den Magen. (nicht: jmdm. auf den Magen schlagen)

– Morphologische Beschränkungen müssen aus der Nennform ersichtlich sein:

Jmdm. ist eine Laus über die Leber gelaufen.

– Die Nennform darf nicht restriktiver als nötig sein:

etwas vom Zaun brechen (etwas = kommunikative Handlung, an der mehr als eine Person (Partei) beteiligt ist)

sich ins Bockshorn jagen lassen (nicht: sich nicht ins Bockshorn jagen lassen)

Man findet in zeitgenössischen Texten und in den Korpora immer wieder Verwendungen von Idiomen, bei denen die Restriktionsbedingungen nicht beachtet sind wie z.B. *Er frisst immer wieder einen Narren an ihr*. Solche Fälle kommen häufig in sprachspielerischen und das heißt in hochgradig markierten Kontexten vor. Der Lexikograph muss hier die Entscheidung treffen, in wie weit er solche Erscheinungen berücksichtigen will. Unser Vorschlag lautet: Markierte Formen sollten nicht als Nennformen erscheinen.

6. Fazit

Die Berücksichtigung des lexikalischen Status von Idiomen ist für deren Lemmatisierung relevant. Dies gilt insbesondere für die Unterscheidung zwischen teil- und voll-idiomatischen Ausdrücken, die sowohl für Wörterbuchbenutzer als auch für Lexikographen durchsichtig und daher leicht zu treffen ist.

Die Teilbarkeit von Idiomen spielt für die Lemmatisierung keine Rolle. Die Berücksichtigung der Teilbarkeit kann sogar dazu führen, dass die Anzahl der Bedeutungspunkte eines Wortes unnötig vermehrt wird. Probleme bei der Zuordnung von Idiomen zu den einzelnen Bedeutungspunkten des Lemmas können dadurch vermieden werden, dass die Idiome am Ende eines Artikels aufgeführt werden.

Für die Auswahl der Nennform sind vor allem obligatorische Beschränkungen hinsichtlich der Besetzung der Argumentstellen des Idioms sowie der Unterschied zwischen Ausdrücken, die Teile des Idioms sind und solchen, die nicht Teile des Idioms sind, ausschlaggebend.

Literatur

- Baranov, Anatolij/Dobrovol'skij, Dmitrij (1999): Idioms from a Cognitive Perspective. In: Moscow State University Bulletin 19. 1/1999, S. 64-75.
- Braasch, Anna (1988): Zur lexikographischen Kodifizierung von Phrasemen in einsprachigen deutschen Wörterbüchern aus der Sicht eines ausländischen Wörterbuchbenutzers. In: Hyldgaard-Jensen, Karl/Zettersten, Arne (Hg.): Proceedings of the Fourth International Symposium on Lexicography. April 20-22, 1988. University of Copenhagen. Lexicographica Series Maior; 26. Tübingen. S. 83-100.
- Burger, Harald (1983): Phraseologie in den Wörterbüchern des heutigen Deutsch. In: Wiegand, Herbert Ernst (Hg.): Studien zur Neuhochdeutschen Lexikographie III. Hildesheim. S. 13-66.
- Burger, Harald (1998): Phraseologie: Eine Einführung am Beispiel des Deutschen. Grundlagen der Germanistik; 36. Berlin.
- Dobrovol'skij, Dmitrij (1995): Kognitive Aspekte der Idiom-Semantik: Studien zum Thesaurus deutscher Idiome. (Eurogermanistik 8). Tübingen.
- Dobrovol'skij, Dmitrij (2000a): Syntaktische Modifizierbarkeit von Idiomem aus lexikographischer Perspektive. In: Heid, Ulrich et al. (Hg.): The Ninth EURALEX International Congress. August 8-12, 2000. Universität Stuttgart. Proceedings. Vol. II. S. 557-568.
- Dobrovol'skij, Dmitrij (2000b): Ist die Semantik von Idiomem nichtkompositionell? In: Beckmann, Susanne/König, Peter-Paul/Wolf, Georg (Hg.): Sprachspiel und Bedeutung: Festschrift für Franz Hundsnurscher zum 65. Geburtstag. Tübingen. S. 113-124.
- DUDEN=DUDEN Das große Wörterbuch der deutschen Sprache in zehn Bänden (1999). 3., völlig neu bearbeitete und erweiterte Auflage. Herausgegeben vom Wissenschaftlichen Rat der Dudenredaktion. Mannheim.
- DUW=DUDEN Deutsches Universalwörterbuch A-Z. (1996) 3. neu bearbeitete und erweiterte Auflage. Auf der Grundlage der neuen amtlichen Rechtschreibregeln. Bearbeitet von Günther Drosdowski und der Dudenredaktion. Mannheim.
- DUDEN/OXFORD=DUDEN OXFORD Großwörterbuch Englisch (1990). Herausgegeben von der Dudenredaktion und Oxford University Press. Mannheim.
- Fleischer, Wolfgang (1997): Phraseologie der deutschen Gegenwartssprache. 2. durchgesehene und ergänzte Auflage. Tübingen.
- Fraser, Bruce (1970): Idioms within a Transformational Grammar. In: Foundations of Language 6, S. 22-42.
- HDG=Handwörterbuch der deutschen Gegenwartssprache in zwei Bänden (1984). Von einem Autorenkollektiv unter der Leitung von Günter Kempcke. Berlin.
- WDG=Klappenbach, Ruth/Steinitz, Wolfgang (Hg.) (1980). Wörterbuch der deutschen Gegenwartssprache (1980). 10., bearbeitete Auflage. Berlin.
- Nunberg, Geoffrey/Sag, Ivan A./Wasow, Thomas (1994): Idioms. In: Language 70, S. 491-538.

Scholze-Stubenrecht, Werner (1988): Phraseologismen im Wörterbuch. In: Harras, Gisela (Hg.): Das Wörterbuch: Artikel und Verweisstrukturen. Jahrbuch 1987 des Instituts für Deutsche Sprache. Düsseldorf. S. 284-302.

Schemann, Hans (1991). Die Phraseologie im zweisprachigen Wörterbuch. In: Hausmann, Franz Josef et al. (Hg.): Wörterbücher: Ein Internationales Handbuch zur Lexikographie. 3. Teilband. Berlin. S. 2789-2794.

Prof. Dr. Gisela Harras
Institut für Deutsche Sprache
Postfach 101621
D-68016 Mannheim
E-Mail: harras@ids-mannheim.de

Kristel Proost
Institut für Deutsche Sprache
Postfach 101621
D-68016 Mannheim
E-Mail: prost@ids-mannheim.de